Deteksi Teks yang Mengandung Keinginan Bunuh Diri pada Media Sosial X Menggunakan Metode Hybrid Deep Learning CNN-BiLSTM

1st Nurlailiyah Salsabilah Valentina
Informatika
Telkom University
Bandung, Indonesia
salsabilavalenn@student.telkomuniversity.ac.id

2nd Erwin Budi Setiawan
Informatika
Telkom University
Bandung, Indonesia
erwinbudisetiawan@telkomuniversity.ac.id

***Abstrak* — Bunuh diri adalah salah satu penyebab kematian terbesar di dunia. Diantara banyak individu yang mengalami tekanan mental, sebagian memilih mengekspresikan perasaan mereka melalui media sosial seperti X. Unggahan-unggahan tersebut sering terdapat tanda-tanda adanya keinginan untuk mengakhiri hidup, yang dapat menjadi peringatan dini jika terdeteksi dengan tepat. Penelitian ini bertujuan mengembangkan sistem yang mampu mendeteksi teks dengan indikasi keinginan bunuh diri. Sistem ini dibangun dengan menggabungkan model deep learning CNN dan BiLSTM, menggunakan TF-IDF sebagai metode ekstraksi fitur dan FastText sebagai metode ekspansi fitur dengan menggunakan corpus similarity sebanyak 64.045 data. Serangkaian skenario dilakukan pada model yang dibangun menggunakan 49.022 data dalam bahasa Inggris yang dikumpulkan dari platform X. Hasil dari pengujian menunjukkan bahwa model hybrid mendapatkan nilai akurasi terbaik. Model hybrid yang mendapatkan nilai akurasi terbaik tersebut adalah CNN-BiLSTM dengan nilai akurasi sebesar 79,21% yang mana mendapatkan kenaikan akurasi dari model baseline sebesar 6,9%. Hasil tersebut menunjukkan bahwa model hybrid dengan kombinasi ekstraksi fitur TF-IDF dan ekspansi fitur FastText mampu mendeteksi indikasi keinginan bunuh diri pada unggahan di sosial media X.**

***Kata kunci*— deteksi bunuh diri, hybrid deep learning, TF-IDF, FastText, X**

I. PENDAHULUAN

Bunuh diri merupakan salah satu masalah kesehatan masyarakat yang serius, dengan angka kematian mencapai sekitar 0,7 juta orang meninggal setiap tahun, terutama pada kalangan muda dan paruh baya [1]. Tindakan ini dipengaruhi oleh beberapa faktor yang dkelompokan menjadi tiga kategori, yaitu faktor kesehatan, faktor lingkungan, dan faktor yang terkait dengan Riwayat pribadi [2]. Faktor tersebut akan memicu adanya perasaan yang tidak dapat diterima yang kemudian memunculkan ide untuk melukai diri sendiri hingga tindakan bunuh diri [3]. Ide melakukan tindakan bunuh diri dapat ditangani oleh ahli medis [4]. Akan tetapi, karena berbagai alasan seperti stigma sosial terhadap masalah mental, sebagian individu menghindari pengobatan medis dan memilih untuk mengekspresikan perasaan mereka melalui *platform* media sosial [5], [6]. Salah satu *platform* media sosial yang digunakan untuk mengekspresikan perasaan adalah Twitter, yang sekarang disebut dengan X.

X adalah *platform* media sosial yang paling umum digunakan untuk membagikan kata-kata, foto, video, dan tautan untuk saling berinteraksi [7]. Beberapa indvidu memanfaatkan anonimitas yang tersedia untuk berbagi kekhawatiran dan ketegangan yang mereka alami, termasuk mengekspresikan keinginan atau niat bunuh diri, mencari mengenai nasihat bunuh diri, bahkan percakapan yang mendorong tindakan bunuh diri [8]. Penggunaan X sebagai sarana menyampaikan perasaan berpotensi memiliki risiko penularan ide bunuh diri kepada individu lain [9]. Oleh karena itu, data pengguna di X dapat digunakan untuk mendeteksi tanda-tanda adanya indikasi keinginan bunuh diri dan melakukan pencegahan terhadap tindakan tersebut.

Dalam mengatasi masalah tersebut, telah dilakukan berbagai penelitian terkait deteksi bunuh diri, beberapa diantaranya menggunakan model gabungan *deep learning Convolutional Neural Network* (CNN) dan *Long-Short Term Memory* (LSTM), yang menghasilkan akurasi sebesar 90,3% [10]. Penelitian lainnya melakukan pengujian menggunakan model *hybrid Convolutional Neural Network* (CNN) dan *Bidirectional Long-Short Term Memory* (BiLSTM) yang mendapatkan hasil akurasi sebesar 94,29% [11]. Selain itu, penelitian lainnya yang juga menggunakan model *hybrid* CNN-BiLSTM mendapatkan nilai akurasi sebesar 95% [12].

Penelitian ini menggunakan model *hybrid deep learning Convolutional Neural Network* (CNN) dan *Bidirectional Long-Short Term Memory* (BiLSTM) untuk mendeteksi teks yang mengandung indikasi bunuh diri. CNN digunakan untuk mengklasifikasikan dan menganalisis teks, sementara BiLSTM digunakan untuk mempertimbangkan konteks menyeluruh dengan memproses informasi dari dua arah [13]. Penggabungan kedua model tersebut dapat meningkatkan akurasi prediksi [14] dalam mendeteksi adanya indikasi keinginan bunuh diri.

II. KAJIAN TEORI

Penelitian terkait deteksi indikasi keinginan bunuh diri diantaranya menerapkan *hybrid deep learning*, seperti yang dilakukan oleh Renjith Shini, dkk [10]. Dengan menggunakan Word2Vec sebagai *word embedding* pada input model CNN-Attention-LSTM dan juga pengoptimalan parameter, diperoleh nilai akurasi sebesar 90,3%. Dikarenakan dataset tidak seimbang, evaluasi model dengan nilai akurasi belum tentu tepat. Sehingga dilakukan evaluasi menggunakan ukuran yang akurat seperti presisi mendapat nilai 91,6%, recall mendapat nilai sebesar 93,7%, dan F1-score mendapat nilai sebesar 92,6%.

Penelitian lainnya dilakukan oleh Mohaiminul Islam Bhuiyan, dkk [11] yang melakukan pengujian *hybrid deep learning* CNN dan BiLSTM dengan menggunakan *word embedding* Word2Vec sebagai ekspansi fiturnya. Penelitian ini melakukan *fine tuning* pada berbagai komponen model untuk meningkatkan kinerja model. Salah satu penyesuaian yang dilakukan adalah teknik regulasi terutama pada L2 *regularization* dan layer *dropout*. Selain itu, digunakan *early stopping* untuk menghindari terjadinya *overfitting*. Dari pengujian yang dilakukan, didapat bahwa model *hybrid* yang dilakukan *fine tuning* mendapatkan peningkatan akurasi dibandingkan model awal yaitu sebesar 92,81%.

Penelitian lainnya terkait deteksi keinginan bunuh diri dilakukan oleh Aldhyani, dkk [12]. Penelitian ini melakukan pengujian pada *hybrid deep learning* CNN dan BiLSTM yang digabungkan dengan ekstraksi fitur TF-IDF dan Word2Vec, serta membandingkannya dengan model XGBoost. Model *hybrid* CNN-BiLSTM mampu menangkap pola linguistik yang kompleks dalam mendeteksi bunuh diri. diketahui bahwa, model *hybrid deep learning* CNN-BiLSTM mendapatkan hasil akurasi terbaik sebesar 95% dan lebih baik dibandingkan dengan model XGBoost yang mendapatkan akurasi terbaik sebesar 91,5%.

Berdasarkan penelitian sebelumnya yang berkaitan dengan deteksi indikasi bunuh diri, dapat dikatakan bahwa model gabungan dari *deep learning* memiliki nilai akurasi yang lebih baik.

III. METODE

Dalam membangun sistem untuk mendeteksi indikasi keinginan bunuh diri pada *platform* X, penelitian ini menggunakan model *hybrid deep learning*. Tahapan yang diusulkan meliputi: data *crawling*, data *labeling*, data *pre-processing*, data *split*, ekstraksi fitur TF-IDF, ekspansi fitur FastText, klasifikasi model, dan evaluasi model yang telah dibangun.

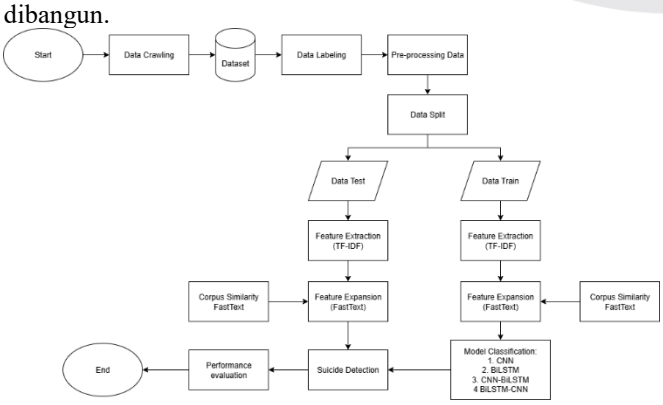

Figure 1. Sistem deteksi bunuh diri

A. Pengumpulan Data

Pengumpulan data melalui proses *crawling* merupakan pengambilan informasi dari berbagai sumber seperti blog, sosial media, atau situs web lainnya. Pada penelitian ini, pengumpulan data bersumber dari *platform* media sosial X berbahasa Inggris. Proses *crawling* dilakukan dengan memanfaatkan API yang disediakan oleh *platform* X. Data yang diambil berisi pesan-pesan yang mengungkapkan perasaan putus asa, kesepian, bahkan perasaan tertekan yang berpotensi memiliki kemungkinan ide bunuh diri. Pesan-pesan tersebut umumnya memiliki berbagai kata kunci, sehingga penelitian ini dipilih kata kunci yang bersifat spesifik dalam mendeteksi indikasi keinginan bunuh diri. Daftar kata kunci yang digunakan dalam proses *crawling* dapat dilihat pada Table 1.

Table 1. Daftar kata kunci

Kata Kunci	Total
Die	10.100
Suicide	15.814
Hopeless	10.003
Depressed	11.003
Kill	2.102
Total	**49,022**

B. Pelabelan data

Setelah pengumpulan data, dilakukan pelabelan agar memudahkan prosess klasifikasi data. Pelabelan dilakukan dengan memberikan nilai 0 untuk teks yang tidak mengindikasikan keinginan bunuh diri, dan nilai 1 untuk teks yang memiliki indikasi keinginan bunuh diri. Penentuan nilai label dilakukan secara berkelompok dengan melibatkan 3 anotator dengan prinsip *majority vote*. Anotator yang terlibat dalam pelabelan adalah anotator yang melakukan penelitian yang sama dengan topik penulis, sehingga hasil dari pelabelan layak untuk dipertanggung jawabkan. Pada Table 2 menunjukkan jumlah data dengan label 0 dan 1. Dapat dilihat bahwa distribusi data tidak seimbang.

Table 2. Jumlah label data

Label	Total	Persentase(%)
0	34.599	70,57
1	14.423	29,42

C. Pre-processing Data

Data hasil *crawling* masih mentah dan tidaak terstruktur [15] sehingga masih memiliki banyak gangguan. Oleh karena itu, diperlukan *preprocessing* untuk mempersiapkan data agar lebih mudah untuk diolah. *Preprocessing* terdiri dari beberapa tahap. Pertama, data *cleaning*, yaitu pembersihan data dari unsur simbol atau karakter yang tidak relevan seperti penghapusan simbol, angka, *hashtag* (#), nama pengguna (@), URL, *emotticon*, spasi berlebih, hapus spasi diawal dan diakhir, serta penghaapusan huruf berulang. Kedua, *case folding* untuk mengubah seluruh teks yang memiliki huruf kapital menjadi huruf kecil. Ketiga, *normalization*, yaitu proses mengubah kata-kata yang tidak baku dalam teks menjadi kata baku. Keempat, *tokenization*, yaitu proses memecah teks menjadi potongan-potongan kata

atau token. Terakhir, *stopwords removal*, yaitu proses menghilangkan kata-kata umum yang kurang bermakna.

D. Ekstraksi Fitur TD-IDF

Ekstraksi fitur adalah proses menghitung bobot setiap kata dalam teks dan mengubahnya menjadi vektor. Tahap ekstraksi fitur ini bertujuan untuk menghasilkan representasi data yang relevan dan mempertahankan informasi penting [16]. Dalam penelitian ini, ekstraksi fitur yang digunakan adalah *Term Frequency-Invers Document Frequency* (TF-IDF). TF-IDF merupakan gabungan dari *Term Frequency* (TF), yang memberi bobot pada kata berdasarkan jumlah kemunculannya, dan *Invers Document Frequency* (IDF) memberikan nilai lebih pada kata-kata yang jarang muncul dalam dokumen dan mengurangi bobot kata-kata yang sering muncul dalam dokumen [17]. Dengan pendekatan ini, kata-kata yang bersifat umum atau sering muncul di hampir semua dokumen akan memiliki bobot yang rendah, berikut ini adalah rumus dari TF-IDF:

$$tf_{td} = \frac{t}{d} \quad (1)$$

$$idf_t = \log\left(\frac{N}{df(t)}\right) \quad (2)$$

$$TF - IDF_{td} = TF \times IDF \quad (3)$$

Pada tf, diukur jumlah frekuensi kata t dalam dokumen d. Sedangkan idf mengukur seberapa penting kata t dalam seluruh dokumen N, yang memberikan bobot lebih tinggi pada kata yang jarang muncul di dokumen lain. Semakin sering kata muncul dalam dokumen dan semakin jarang kata tersebut muncul di seluruh koleksi dokumen, maka nilai TF-IDF akan semakin tinggi.

Pada ekstraksi fitur, TF-IDF dapat dikombinasikan dengan pendekatan N-Gram, seperti Unigram, Bigran, dan Trigram [18]. Pendekatan ini juga dapat digabungkan dengan beberapa jenis N-Gram untuk memperoleh representasi fitur yang lebih kaya [19]. Penggunaan N-Gram memungkinkan model menangkap hubungan antar kata yang lebih kompleks dan konteks kata yang luas. Penelitian ini menggunakan lima jenis N-Gram, yaitu: Unigram, Bigram, Trigram, Uni-Bigram, dan Uni-Trigram. Unigram merepresentasikan satu kata, Bigram merepresentasikan dua kata berturut-turut, Trigram merepresentasikan tiga kata berturut-turut, Uni-Bigram merepresentasikan satu hingga dua kata, dan Uni-Trigram merepresentasikan satu hingga tiga kata.

E. Ekspansi Fitur FastText

Ekspansi fitur yang digunakan dalam penelitian ini adalah *word embedding* menggunakan FastText. FastText merupakan *library* yang dikembangkan Facebook untuk pembelajaran representasi kata dan klasifikasi kalimat. Teknik ini memperkaya representasi vektor dengan cara memecah kata-kata yang tidak dikenali dalam *corpus* [20]. Setiap kata direpresentasikan sebagai karakter N-Gram, sehingga membantu menangkap arti kata-kata yang pendek [21].

F. Model Klasifikasi

Setelah proses ekstraksi fitur, hasil vektor yang didapat akan digunakan sebagai input untuk model *hybrid deep learning Convolutional Neural Network* (CNN) dan *Bidirectional Long-Short Term Memory* (BiLSTM). CNN memanfaatkan *convolutional layer* yang dapat mengurangi beban komputasi [22]. *Convolutional layer* terbentuk dari beberapa gabungan lapisan konvolusi, lapisan *pooling*, dan lapisan *fully connected* yang digunakan dalam klasifikasi citra [23]. Lapisan konvolusi digunakan untuk ekstraksi fitur lokal, kemudian dilanjutkan pada lapisan *pooling* yang mengurangi dimensi data dan meminimalkan kompleksitas. Selanjutnya, pada tahap *fully connected layer*, keputusan akhir dibuat berdasarkan informasi yang telah diproses dan diekstraksi sebelumnya. Pada Figure 2 merupakan gambar arsitektur model CNN.

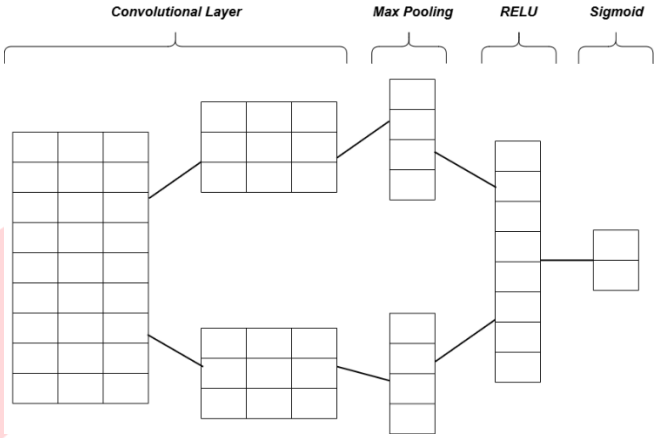

Figure 2. Arsitektur CNN

Bidirectional Long-Short Term Memory (BiLSTM) merupakan variasi LSTM yang terdiri dari dua lapisan, yaitu *forward pass* dan *backward pass*, yang memungkinkan model untuk memperoleh informasi baik sebelum maupun sesudah titik data [24]. Kombinasi dari kedua lapisan ini memungkinkan BiLSTM untuk lebih memahami konteks dan makna dalam urutan data, karena dapat menangkap informasi dari kedua arah, baik masa lalu maupun masa depan [25]. Dengan begitu, BiLSTM mampu memberikan hasil yang lebih akurat. Pada Figure 3 adalah gambar dari arsitektur model BiLSTM.

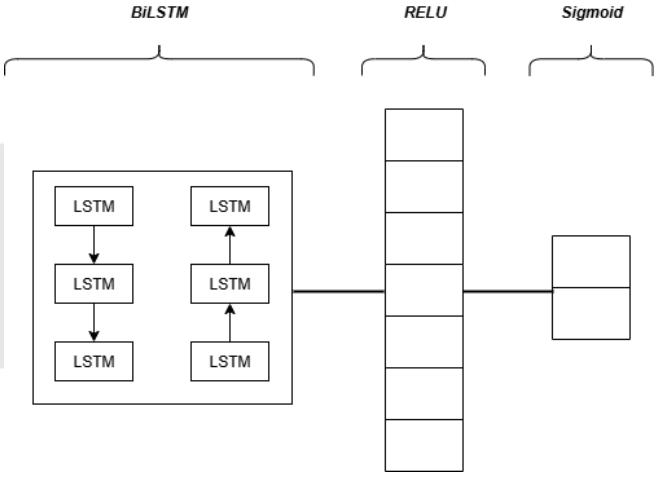

Figure 3. Arsitektur BiLSTM

Penggabungan CNN dan BiLSTM dalam model *hybrid* ini memungkinkan untuk memanfaatkan keunggulan kedua arsitektur. CNN digunakan untuk mengekstraksi pola lokal dan BiLSTM digunakan untuk menangkap depedensi sekuensial dua arah, sehingga meningkatkan pemahaman konteks dan akurasi dalam tugas-tugas pemrosesan bahasa alami seperti klasifikasi teks [26]. Dengan memanfaatkan model *hybrid* ini, dapat secara efektif mengenali fitur lokal

dan global dalam data, yang sangat penting untuk menangani kompleksitas struktur teks yang bervariasi. Figure 4 merupakan arsitektur dari mode *hybrid* CNN-BiLSTM.

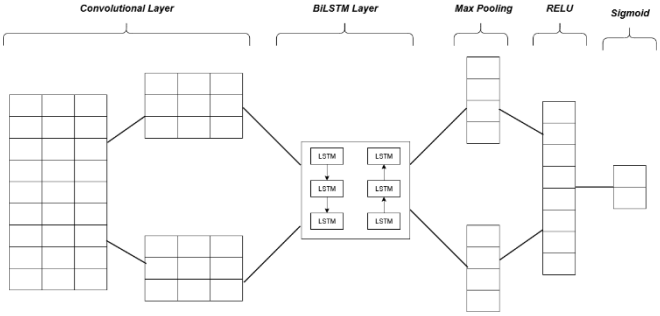

Figure 4. Arsitektur CNN-BiLSTM

G. Evaluasi Performa

Tahap evaluasi performa *hybrid deep learning* dilakukan dengan menggunakan *confusion matrix*. Menurut penelitian Maimi Herawati, dkk [27], *confusion matrix* digunakan sebagai metode dalam mengukur model klasifikasi. Terdapat empat kombinasi yang menggambarkan hasil klasifikasi yaitu *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN). Hasil dari *confusion matrix* digunakan untuk menghitung nilai evaluasi seberapa baik model klasifikasi yang terdiri dari akurasi, presisi, *recall*, dan F1-score. Akurasi mengukur seberapa sering model membuat prediksi yang benar dari seluruh jumlah prediksi, presisi mengukur ketepatan prediksi positif dibandingkan seluruh prediksi positif yang diberikan model, sedangkan *recall* mengukur kemampuan model dalam menemukan semua data positif. Untuk menghitung nilai *confusion matrix* dapat menggunakan pola berikut:

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

$$Presisi = \frac{TP}{TP + FP} \quad (5)$$

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

$$F1 - Score = 2x\frac{presisi\ x\ recall}{presisi + recall} \quad (7)$$

IV. HASIL DAN PEMBAHASAN

Pada penelitian ini, dilakukan empat skenario pengujian terhadap empat model klasifikasi, yaitu CNN, BiLSTM, CNN-BiLSTM, dan BiLSTM-CNN. Skenario pertama bertujuan untuk mencari rasio pembagian data terbaik pada model *baseline* CNN dan BiLSTM. Skenario kedua menerapkan rasio pembagian data dan menguji jumlah maksimal fitur TF-IDF. Skenario ketiga menguji penerapan N-Gram pada TF-IDF. Dan skenario keempat mengkombinasikan model CNN dan BiLSTM dari hasil skenario sebelumnya dengan ekspansi fitur FastText.

A. Hasil Pengujian

Skenario pertama bertujuan untuk menentukan rasio pembagian data terbaik pada model *baseline* menggunakan ekstraksi fitur TF-IDF unigram. Parameter yang digunakan untuk model CNN adalah *filter* 128 dengan *kernel size* 5, sementara untuk BiLSTM menggunakan *dense layer* 64. Kedua model ini menggunakan *batch size* 32 dan *epoch* 5. Rasio pembagian data yang diterapkan adalah 70:30, 80:20, dan 90:10. Dari Table 5, hasil skenario pertama menunjukkan nilai akurasi terbaik model CNN dan gabungan CNN-BiLSTM pada rasio pembagian data 80:20. Sedangkan untuk model BiLSTM dan gabungan BiLSTM-CNN mendapatkan nilai akurasi terbaik pada pembagian data 70:30. Model CNN mendapatkan akurasi sebesar 74,09%, model BiLSTM mendapatkan akurasi 73,05%, model gabungan CNN-BiLSTM mendapatkan akurasi 79,18%, dan model gabungan BiLSTM-CNN mendapatkan akurasi 74,23%. Hasil skenario pembagian data terbaik akan digunakan untuk skenario selanjutnya.

Table 3. Skenario pertama (*Baseline*)

Accuracy (%)	Rasio		
	70:30	**80:20**	**90:10**
CNN	73,69	**74,09**	73,10
BiLSTM	**73,05**	71,86	70,91
CNN-BiLSTM	79,07	**79,18**	78,78
BiLSTM-CNN	**74,23**	72,27	71,88

Skenario kedua adalah menerapkan hasil pembagian data terbaik dengan vektor fitur 2500, 5000, 7500, dan 10000. Table 6 menunjukkan hasil akurasi terbaik menggunakan vektor fitur 5000 dengan model CNN mendapatkan akurasi 74,09%, BiLSTM 73,05%, CNN-BiLSTM 79,18%, dan BiLSTM-CNN 74,23%. Hasil tersebut akan digunakan untuk skenario pengujian selanjutnya.

Table 4. Skenario kedua (Max Feature)

Accuracy (%)	Max Feature			
	2500	**5000**	**7500**	**10000**
CNN	72,94	**74,09**	72,50	73,86
BiLSTM	69,40	**73,05**	71,60	70,29
CNN-BiLSTM	75,62	**79,18**	78,83	75,55
BiLSTM-CNN	68,13	**74,23**	72,15	71,77

Skenario ketiga bertujuan untuk menerapkan kombinasi parameter N-Gram, yaitu unigram, brigram, trigram, uni-bigram, uni-trigram. Table 7 menunjukkan bahwa penerapan parameter N-Gram menghasilkan hasil terbaik pada model CNN dan gabungan BiLSTM-CNN dengan parameter bigram. Sedangkan model gabungan BiLSTM dan CNN-BiLSTM mendapatkan hasil terbaik dengan menggunakan parameter N-Gram unigram. Hasil uji ini akan digunakan untuk skenario pengujian selanjutnya.

Table 5. Skenario ketiga (N-Gram)

N-Gram	Accuracy (%)			
	CNN	**BiLSTM**	**CNN-BiLSTM**	**BiLSTM-CNN**
Unigram	74,09	**73,05**	**79,18**	74,23
Bigram	**74,92**	58,58	78,68	**75,14**
Trigram	73,66	64,44	76,53	74,08
Uni-Bigram	73,09	69,87	77,27	72,29
Uni-Trigram	72,53	67,69	73,38	64,09

Skenario keempat menerapkan ekspansi fitur FastText sebagai *word embedding* pada klasifikasi model CNN, BiLSTM, CNN-BiLSTM, dan BiLSTM-CNN dari hasil skenario sebelumnya. Pengujian dilakukan dengan memilih

fitur-fitur teratas dari *corpus*. Fitur-fitur teratas yang digunakan adalah top 1, top 5, top 10, top 15 dan top 20. Pada model *hybrid* CNN-BiLSTM, ukuran padding menggunakan same tetapi model lainnya menggunakan padding default. Table 8 menunjukkan bahwa model *hybrid* mendapatkan hasil akurasi yang baik.

Rank	Accuracy (%)			
	CNN	BiLSTM	CNN-BiLSTM	BiLSTM-CNN
Top 1	75,21	61,28	**79,21**	**75,98**
Top 5	**75,73**	72,80	77,36	75,14
Top 10	75,40	74,46	72,87	64,51
Top 15	75,58	**75,00**	74,51	65,35
Top 20	-	71,73	-	-

B. Analisis Hasil Pnegujian

Setelah rangkaian pengujian dilakukan untuk menentukan model *baseline* terbaik melalui pembagian data, vektor fitur, kombinasi nilai N-Gram dengan TF-IDF, dan top rank terbaik dalam implementasi ekspansi fitur menggunakan FastText.

Berdasarkan hasil pengujian yang telah dilakukan, diperoleh bahwa skenario pertama model *baseline* terbaik pada pembagian data 80:20 adalah model CNN yang mendapatkan nilai akurasi sebesar 74,09% dan model gabungan CNN-BiLSTM yang mendapatkan hasil akurasi sebesar 79,18%. Dan pada pembagian data 70:30 adalah model BiLSTM dengan nilai akurasi sebesar 73,05% dan model gabungan BiLSTM-CNN dengan nilai akurasi sebesar 74,23%.

Skenario kedua, jumlah vektor fitur terbaik didapat pada nilai 5000. Pada skenario ini tidak mendapatkan peningkatan akurasi dikarenakan jumlah vektor fitur yang digunakan sama dengan pengujian skenario pertama. Model CNN mendapatkan nilai akurasi sebesar 74,09%, model BiLSTM memperoleh hasil akurasi sebesar 73,05%, model gabungan CNN-BiLSTM memperoleh hasil akurasi sebesar 79,18%, dan model gabungan BiLSTM-CNN mendapatkan hasil akurasi sebesar 74,23%.

Pada skenario ketiga, beberapa model mendapatkan peningkatan akurasi dengan menggabungkan TF-IDF dengan bigram. Model yang mendapatkan peningkatan akurasi adalah model CNN yang memperoleh hasil akurasi 74,92% dengan peningkatan akurasi sebesar 0,83%, dan model gabungan BiLSTM-CNN yang memperoleh nilai akurasi sebesar 75,14% dengan peningkatan akurasi sebesar 0,91%.

Pada skenario keempat, pengimplementasian ekspansi fitur pada setiap model. Model CNN memperoleh akurasi terbaik sebesar 75,73% dengan peningkatan akurasi dari model *baseline* sebesar 1,64% pada peringkat korpus Top 5. Model BiLSTM memperoleh nilai akurasi terbaik sebesar 75,00% dengan pengingkatan akurasi dari model *baseline* sebesar 1,95% pada peringkat korpus Top 15. Sedangkan pada model *hybrid* mendapatkan akurasi terbaik pada peringkat korpus Top 1 dengan model CNN-BiLSTM mendapatkan nilai akurasi sebesar 79,21% dan model BiLSTM-CNN memperoleh hasil akurasi sebesar 75,98%.

Peningkatan akurasi dari semua skenario dapat dilihat pada Figure 5.

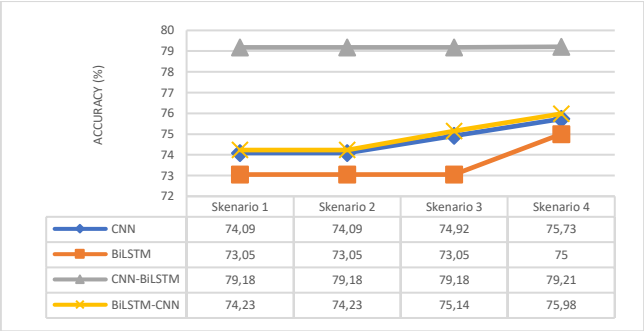

Figure 5. Perbandingan hasil semua skenario

Hasil penelitian menunjukkan bahwa model *hybrid deep learning* CNN-BiLSTM dengan ekstraksi fitur TF-IDF serta ekspansi fitur menggunakan FastText menghasilkan nilai akurasi tertinggi dalam mendeteksi indikasi keinginan bunuh diri. Karena permasalahan bunuh diri sangat penting, penilaian model tidak hanya berfokus pada akurasi saja, tetapi juga memperhatikan presisi, recall dan F1-score guna memastikan kemampuan deteksi yang optimal seperti yang ditunjukkan pada Table 9.

Table 6. Nilai *confusion matrix*

Model	Akurasi	Presisi	Recall	F1-Score
CNN	75,73	65,79	36,43	46,70
BiLSTM	75,00	67,55	31,45	42,19
CNN-BiLSTM	79,21	87,68	33,87	48,78
BiLSTM-CNN	75,98	69,81	33,50	44,75

Dapat dilihat bahwa nilai akurasi sudah mendapatkan hasil yang cukup baik. Akan tetapi, nilai *confusion matrix*, seperti *recall*, mendapatkan nilai yang rendah yang menyatakan bahwa model belum cukup baik dalam mendeteksi indikasi keinginan bunuh diri. Kondisi ini kemungkinan disebabkan oleh ketidakseimbangan data. Untuk mengatasi hal tersebut, penelitian ini telah melakukan beberapa upaya, antara lain menerapkan teknik *Synthetic Minority Over-sampling Technique* (SMOTE), melakukan penyesuaian parameter, serta mengganti dan mengoptimalkan *optimizer*. Namun, berbagai langkah tersebut belum mampu memberikan peningkatan yang signifikan pada nilai *confusion matrix*.

V. KESIMPULAN

Penelitian ini melakukan deteksi indikasi keinginan bunuh diri pada unggahan teks berbahasa Inggris pada X menggunakan empat model klasifikasi, yaitu CNN, BiLSTM, *hybrid* CNN-BiLSTM, dan *hybrid* BiLSTM-CNN. Model-model ini menggunakan TF-IDF dalam ekstraksi fitur serta *word embedding* FastText untuk pembuatan *corpus* dan ekspansi fitur. Hasil pengujian menunjukkan bahwa model *hybrid* memiliki akurasi lebih tinggi dibanding model tunggal. Penerapan kombinasi TF-IDF dengan N-Gram serta ekspansi fitur menggunakan FastText berhasil memperkaya representasi kata dan meningkatkan akurasi. Model *hybrid* CNN-BiLSTM memberikan hasil terbaik dengan akurasi sebesar 79,21%, lalu model *hybrid* BiLSTM-CNN dengan akurasi sebesar 75,98%. Hal ini membuktikan bahwa kombinasi TF-IDF dan FastText dapat meningkatkan performa model dalam deteksi indikasi keinginan bunuh diri.

Meskipun demikian, penelitian ini memiliki beberapa keterbatasan. Pertama, proses pelabelan tidak melibatkan ahli dibidang tersebut. Kedua, meskipun telah melakukan penyesuaian *hyperparameter*, penanganan data yang tidak seimbang menggunakan SMOTE, serta penyesuaian *optimizer*, peningkatan nilai metrik lain pada *confusion matrix* seperti presisi, *recall*, dan F1-score belum signifikan.

Untuk penelitian selanjutnya, disarankan untuk melibatkan anotator ahli untuk memastikan kualitas label. Menggunakan metode penanganan data yang tidak seimbang lainnya seperti ADASYN dan lain sebagainya, serta mempertimbangkan model berbasis *transformer* seperti BERT atau RoBERTa untuk meningkatkan kemampuan pemahaman konteks.

REFERENSI

[1] R. Haque, N. Islam, M. Islam, and M. M. Ahsan, "A Comparative Analysis on Suicidal Ideation Detection Using NLP, Machine, and Deep Learning," *Technologies (Basel)*, vol. 10, no. 3, 2022, doi: 10.3390/technologies10030057.

[2] E. Yeskuatov, S. L. Chua, and L. K. Foo, "Leveraging Reddit for Suicidal Ideation Detection: A Review of Machine Learning and Natural Language Processing Techniques," 2022. doi: 10.3390/ijerph191610347.

[3] Ririn Nasriati, Filia Icha, and Metti Verawati, "DETEKSI KEGAWATDARURATAN PSIKIATRI PADA REMAJA DI PONOROGO," *Nursing Sciences Journal*, vol. 7, no. 2, 2023, doi: 10.30737/nsj.v7i2.5060.

[4] J. Liu, M. Shi, and H. Jiang, "Detecting Suicidal Ideation in Social Media: An Ensemble Method Based on Feature Fusion," *Int J Environ Res Public Health*, vol. 19, no. 13, 2022, doi: 10.3390/ijerph19138197.

[5] P. Burnap, G. Colombo, R. Amery, A. Hodorog, and J. Scourfield, "Multi-class machine classification of suicide-related communication on Twitter," *Online Soc Netw Media*, vol. 2, 2017, doi: 10.1016/j.osnem.2017.08.001.

[6] K. S. Seby, M. Elamparithi, and V. Anuratha, "Suicidal Ideation Detection from social media: A Detailed Review of Machine Learning and Deep Learning," *International Journal of Membrane Science and Technology*, vol. 10, no. 2, 2023, doi: 10.15379/ijmst.v10i2.2958.

[7] E. L. Unruh-Dawes, L. M. Smith, C. P. Krug Marks, and T. T. Wells, "Differing Relationships Between Instagram and Twitter on Suicidal Thinking: The Importance of Interpersonal Factors," *Social Media and Society*, vol. 8, no. 1, Feb. 2022, doi: 10.1177/20563051221077027.

[8] M. Chatterjee, P. Kumar, P. Samanta, and D. Sarkar, "Suicide ideation detection from online social media: A multi-modal feature based technique," *International Journal of Information Management Data Insights*, vol. 2, no. 2, 2022, doi: 10.1016/j.jjimei.2022.100103.

[9] J. Luo, J. Du, C. Tao, H. Xu, and Y. Zhang, "Exploring temporal suicidal behavior patterns on social media: Insight from Twitter analytics," *Health Informatics J*, vol. 26, no. 2, 2020, doi: 10.1177/1460458219832043.

[10] S. Renjith, A. Abraham, S. B. Jyothi, L. Chandran, and J. Thomson, "An ensemble deep learning technique for detecting suicidal ideation from posts in social media platforms," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 10, 2022, doi: 10.1016/j.jksuci.2021.11.010.

[11] M. I. Bhuiyan, N. S. Kamarudin, and N. H. Ismail, "Enhanced Suicidal Ideation Detection from Social Media Using a CNN-BiLSTM Hybrid Model," Jan. 2025, [Online]. Available: http://arxiv.org/abs/2501.11094

[12] T. H. H. Aldhyani, S. N. Alsubari, A. S. Alshebami, H. Alkahtani, and Z. A. T. Ahmed, "Detecting and Analyzing Suicidal Ideation on Social Media Using Deep Learning and Machine Learning Models," *Int J Environ Res Public Health*, vol. 19, no. 19, 2022, doi: 10.3390/ijerph191912635.

[13] L. Xiaoyan, R. C. Raga, and S. Xuemei, "GloVe-CNN-BiLSTM Model for Sentiment Analysis on Text Reviews," *J Sens*, vol. 2022, 2022, doi: 10.1155/2022/7212366.

[14] M. Marjani, M. Mahdianpari, and F. Mohammadimanesh, "CNN-BiLSTM: A Novel Deep Learning Model for Near-Real-Time Daily Wildfire Spread Prediction," *Remote Sens (Basel)*, vol. 16, no. 8, Apr. 2024, doi: 10.3390/rs16081467.

[15] A. Gasparetto, M. Marcuzzo, A. Zangari, and A. Albarelli, " Survey on Text Classification Algorithms: From Text to Predictions," *Information (Switzerland)*, vol. 13, no. 2, 2022, doi: 10.3390/info13020083.

[16] C. Janiesch, P. Zschech, and K. Heinrich, "Machine learning and deep learning", doi: 10.1007/s12525-021-00475-2/Published.

[17] K. Kowsari, K. J. Meimandi, M. Heidarysafa, S. Mendu, L. Barnes, and D. Brown, "Text classification algorithms: A survey," 2019. doi: 10.3390/info10040150.

[18] Sutriawan, Muljono, Khairunnisa, Z. Alamin, T. A. Lorosae, and S. Ramadhan, "Improving Performance Sentiment Movie Review Classification Using Hybrid Feature TFIDF, N-Gram, Information Gain and Support Vector Machine," *Mathematical Modelling of Engineering Problems*, vol. 11, no. 2, 2024, doi: 10.18280/mmep.110209.

[19] "Penerapan Algoritma Support Vector Machine (SVM) dengan TF-IDF N-Gram untuk Text Classification _ Arifin _ STRING (Satuan Tulisan Riset dan Inovasi Teknologi)".

[20] R. Adipradana, B. P. Nayoga, R. Suryadi, and D. Suhartono, "Hoax analyzer for indonesian news using rnns with fasttext and glove embeddings," *Bulletin of Electrical Engineering and Informatics*, vol. 10, no. 4, 2021, doi: 10.11591/eei.v10i4.2956.

[21] A. Nurdin, B. Anggo Seno Aji, A. Bustamin, and Z. Abidin, "PERBANDINGAN KINERJA WORD EMBEDDING WORD2VEC, GLOVE, DAN

FASTTEXT PADA KLASIFIKASI TEKS," *Jurnal Tekno Kompak*, vol. 14, no. 2, 2020, doi: 10.33365/jtk.v14i2.732.

[22] Yudi Widhiyasana, Transmissia Semiawan, Ilham Gibran Achmad Mudzakir, and Muhammad Randi Noor, "Penerapan Convolutional Long Short-Term Memory untuk Klasifikasi Teks Berita Bahasa Indonesia," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, vol. 10, no. 4, 2021, doi: 10.22146/jnteti.v10i4.2438.

[23] A. Mulyanto, E. Susanti, F. Rossi, W. Wajiran, and R. I. Borman, "Penerapan Convolutional Neural Network (CNN) pada Pengenalan Aksara Lampung Berbasis Optical Character Recognition (OCR)," *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, vol. 7, no. 1, 2021, doi: 10.26418/jp.v7i1.44133.

[24] D. Oleh, M. Dzaki, and A. Nasir, "OPTIMASI MODEL BILSTM UNTUK ANALISIS SENTIMEN ULASAN FILM MENGGUNAKAN HYPERPARAMETER TUNING RANDOM SEARCH," 2023.

[25] A. Rofiqul Muslikh, I. Akbar, D. Rosal Ignatius Moses Setiadi, and H. Md Mehedul Islam, "Multi-label Classification of Indonesian Al-Quran Translation based CNN, BiLSTM, and FastText." [Online]. Available: https://quran.kemenag.go.id.

[26] A. Mas'ud, B. K. Triwijoyo, and D. Priyanto, "Prediksi Gender Berdasarkan Nama Menggunakan Kombinasi Model IndoBERT, Convolutional Neural Network (CNN) dan Bidirectional Long Short-Term Memory (BiLSTM)," *JTIM : Jurnal Teknologi Informasi dan Multimedia*, vol. 7, no. 3, pp. 448–460, Jun. 2025, doi: 10.35746/jtim.v7i3.736.

[27] M. Herawati, "Prediksi Transaksi Minat Pembelian Online Menggunakan Kombinasi CNN Conv1D dan BiLSTM," *BiLSTM). Journal Cerita: Creative Education of Research in Information Technology and Artificial Informatics*, vol. 11, no. 1, pp. 120–128, 2025, doi: 10.33050/cerita.v11i1.3702.